R. Peled *et al.* (2017) "On the Wegner Orbital Model,"
International Mathematics Research Notices, Vol. 2019, No. 4, pp. 1030–1058
Advance Access Publication July 11, 2017
doi: 10.1093/imrn/rnx145

On the Wegner Orbital Model

Ron Peled[1], Jeffrey Schenker[2], Mira Shamis[3,4], and Sasha Sodin[1,4,*]
[1]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel, [2]Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA, [3]Department of Mathematics, The Weizmann Institute of Science, Rehovot 7610001, Israel, and [4]School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK

[*]*Correspondence to be sent to: e-mail: sashas***1***@post.tau.ac.il*

The Wegner orbital model is a class of random operators introduced by Wegner to model the motion of a quantum particle with many internal degrees of freedom (orbitals) in a disordered medium. We consider the case when the matrix potential is Gaussian, and prove three results: localisation at strong disorder, a Wegner-type estimate on the mean density of eigenvalues, and a Minami-type estimate on the probability of having multiple eigenvalues in a short interval. The last two results are proved in the more general setting of deformed block-Gaussian matrices, which includes a class of Gaussian band matrices as a special case. Emphasis is placed on the dependence of the bounds on the number of orbitals. As an additional application, we improve the upper bound on the localisation length for one-dimensional Gaussian band matrices.

1 Statement of Results

The current investigation is motivated by the work of Wegner [42] and its continuation by Schäfer and Wegner [35] and Oppermann and Wegner [30] on the motion of a quantum particle with many ($N \gg 1$) internal degrees of freedom in a disordered medium.

Received January 17, 2017; Revised January 17, 2017; Accepted May 12, 2017
Communicated by Prof. Kurt Johansson

The Hamiltonian H of the quantum particle acts on a dense subset of $\ell_2(\mathbb{Z}^d \to \mathbb{C}^N)$, the space of square-integrable functions from $\mathbb{Z}^d$ to $\mathbb{C}^N$, via

$$(H\psi)(x) = V(x)\psi(x) + \sum_{y\,:\,y\sim x} W(x,y)\psi(y)\ , \quad x \in \mathbb{Z}^d\ , \tag{1.1}$$

where the potential entries $V(x)$ are $N \times N$ Hermitian matrices, and the hopping terms $W(x,y)$ are $N \times N$ matrices with the Hermitian constraint $W(y,x) = W(x,y)^*$. Following [42], we take the potential entries and hopping terms random and assume them to be independent up to the Hermitian constraint, meaning that

$$\left\{V(x) \mid x \in \mathbb{Z}^d\right\} \bigcup \left\{W(x,y) \mid x,y \in \mathbb{Z}^d, x \text{ has even sum of coordinates and } x \sim y\right\}$$

are jointly independent. We assume that either the distribution of each $V(x)$ is given by the Gaussian Orthogonal Ensemble (GOE), and that the matrices $W(x,y)$ are real (orthogonal case), or that the distribution of each $V(x)$ is given by the Gaussian Unitary Ensemble (GUE) (unitary case). Here, the probability density of the GOE with respect to the Lebesgue measure on real symmetric matrices is proportional to $\exp\left\{-\frac{N}{4}\operatorname{tr} V^2\right\}$, and the probability density of the GUE with respect to the Lebesgue measure on Hermitian matrices is proportional to $\exp\left\{-\frac{N}{2}\operatorname{tr} V^2\right\}$.

Special cases of the model (1.1) include the block Anderson model and the Wegner orbital model in their orthogonal and unitary versions, given by

$$\begin{array}{c} \text{block} \\ \text{Anderson} \end{array} \begin{cases} (H^{\mathrm{bA},\mathbb{R}}\psi)(x) = V^{\mathrm{GOE}}(x)\psi(x) + g\sum_{y\sim x}(\psi(x) - \psi(y)), \\ (H^{\mathrm{bA},\mathbb{C}}\psi)(x) = V^{\mathrm{GUE}}(x)\psi(x) + g\sum_{y\sim x}(\psi(x) - \psi(y)), \end{cases}$$
$$\begin{array}{c} \text{Wegner} \\ \text{orbital} \end{array} \begin{cases} (H^{\mathrm{Weg},\mathbb{R}}\psi)(x) = V^{\mathrm{GOE}}(x)\psi(x) + g\sum_{y\sim x} W^{\mathbb{R}}(x,y)\psi(y), \\ (H^{\mathrm{Weg},\mathbb{C}}\psi)(x) = V^{\mathrm{GUE}}(x)\psi(x) + g\sum_{y\sim x} W^{\mathbb{C}}(x,y)\psi(y), \end{cases} \tag{1.2}$$

where $W^{\mathbb{R}}(x,y)$ has independent real Gaussian $N_{\mathbb{R}}(0, 1/N)$ entries, $W^{\mathbb{C}}(x,y)$ has independent complex Gaussian $N_{\mathbb{C}}(0, 1/N)$ entries, $g > 0$ is a coupling constant, and superscripts indicate the symmetry class of the potential matrices. The block Anderson model is a generalisation of the Anderson model [8] with Gaussian disorder (which is recovered when $N = 1$), whereas the Wegner orbital model is invariant in distribution under local gauge transformations, that is, conjugation by $\mathcal{U}$ of the form

$$(\mathcal{U}\psi)(x) = \mathcal{U}(x)\psi(x)\ , \quad \text{where} \quad \mathcal{U}(x) \in \begin{cases} O_N\ , & H^{\mathrm{Weg},\mathbb{R}} \\ U_N\ , & H^{\mathrm{Weg},\mathbb{C}} \end{cases}, \quad x \in \mathbb{Z}^d\ .$$

Our results pertain to three topics: localisation at strong disorder in arbitrary dimension (Theorem 1), estimates on the density of states (Wegner estimates, Theorem 2), and on the probabilities of multiple eigenvalues in a short interval (Minami estimates, Theorem 3). The latter two results are proved in greater generality, for deformed block-Gaussian matrices, and are also applicable to Gaussian band matrices. The common theme is the strive for the sharp dependence on the number N of orbitals (internal degrees of freedom). As an additional application, we improve the upper bound from [36] on the localisation length of one-dimensional Gaussian band matrices (Theorem 4).

Strong disorder localisation.

The Anderson model in dimension $d \geq 3$ is conjectured to exhibit a spectral phase transition between a localisation (insulator) regime and a delocalisation (conductor) regime. In particular, there should exist a threshold $g_0(d)$ such that for $g < g_0(d)$ the spectrum is pure point, whereas for $g > g_0(d)$ it has an absolutely continuous component. So far only the localisation side of the transition has been established mathematically. Two methods of proof are now available: the multi-scale analysis of Fröhlich and Spencer [24] and the fractional moment method of Aizenman and Molchanov [3].

A phase transition similar to that of the Anderson model is conjectured to occur, in dimension $d \geq 3$, for the orbital models (1.2), with the threshold $g_0(d,N)$ depending on the dimension and the number of orbitals. The first subject of the current article is the dependence of the threshold $g_0(d,N)$ on the number of orbitals N. On the physical level of rigour, this question was settled already in the original papers [30, 35, 42]. The arguments provided there indicate that, for $d \geq 3$,

$$g_0(d,N) \sim \left\{C(d)\sqrt{N}\right\}^{-1} \quad \text{as } N \to \infty \,. \tag{1.3}$$

Two heuristic arguments are discussed in Section 4.

Our first result, Theorem 1 below, is a mathematically rigorous confirmation to one direction of (1.3), the localisation side. The result is stated for the general model (1.1). When specialised to the models (1.2), it asserts that, for g below the threshold (1.3), the matrix elements of the resolvent decay exponentially in the distance from the diagonal. The formal statement is in terms of finite-volume restrictions: denote by

$$P_\Lambda : \ell_2(\mathbb{Z}^d \to \mathbb{C}^N) \to \ell_2(\Lambda \to \mathbb{C}^N)$$

the coordinate projection to a finite volume $\Lambda \subset \mathbb{Z}^d$ and, for an operator H of the form (1.1), let

$$H_\Lambda := P_\Lambda H P_\Lambda^* \tag{1.4}$$

be the restriction of H to Λ. Let $\|x-y\|_1$ be the graph distance between $x, y \in \mathbb{Z}^d$, let $\|v\|$ be the ℓ_2 norm of a vector $v \in \mathbb{C}^N$ and let $\|W\|_{\mathrm{op}}$ be the operator norm of a matrix W.

Theorem 1. There exists a constant $C > 0$ such that the following holds. Let $0 < s < 1$, let H be as in (1.1) in either the orthogonal case or the unitary case and suppose that

$$g_{\mathrm{eff}} := \sup_{x,y} \left\{ \mathbb{E}\|W(x,y)\|_{\mathrm{op}}^s \right\}^{\frac{1}{s}} < \left\{ \frac{1-s}{Cd} \right\}^{\frac{1}{s}} \frac{1}{\sqrt{N}} \,. \tag{1.5}$$

Then for any finite $\Lambda \subset \mathbb{Z}^d$, $x, y \in \Lambda$, $\lambda \in \mathbb{R}$, and $v \in \mathbb{C}^N$:

$$\mathbb{E}\|(H_\Lambda - \lambda)^{-1}(x,y)\, v\|^s \le \frac{CN^{s/2}}{1-s} \left(\frac{Cd(g_{\mathrm{eff}}\sqrt{N})^s}{1-s} \right)^{\|x-y\|_1} \|v\|^s \,. \tag{1.6}$$

□

In the left-hand side of (1.6), we first take the matrix inverse, then extract an $N \times N$ block, and then multiply by a vector, or formally:

$$(H_\Lambda - \lambda)^{-1}(x,y)\, v = P_{\{x\}}(H_\Lambda - \lambda)^{-1} P_{\{y\}}^* v \,.$$

Also observe that the assumption (1.5) guarantees that $\frac{Cd(g_{\mathrm{eff}}\sqrt{N})^s}{1-s} < 1$, and thus the right-hand side of (1.6) indeed decays exponentially in $\|x-y\|_1$.

Remark 1.1. Theorem 1 applies to the models (1.2) and yields the conclusion (1.6), where g_{eff} is replaced with g. For $s = 1 - \log^{-1}(d+2)$, the assumption (1.5) is implied by

$$g < \left\{ Cd \log(d+2)\sqrt{N} \right\}^{-1} \tag{1.7}$$

(where the constant may differ from that of (1.5)). □

Remark 1.2. Methods have been developed to pass from decay estimates for the resolvent in finite volume to other signatures of Anderson localisation, in particular, pure point spectrum and dynamical localisation in infinite volume. We refer in particular to the eigenfunction correlator method introduced by Aizenman [2]; see further [5, Theorem A.1]. Such methods can also be applied in our setup. □

The main feature of Theorem 1 is the dependence on the number of orbitals, N, which, for the models (1.2), is conjecturally sharp in dimension $d \geq 3$. For comparison, localisation for

$$g < \left\{C(d)N^{3/2}\right\}^{-1}$$

follows from the general theorems pertaining to variants of the Anderson model (proved either by the method of [24], or of [3]) and does not require the additional arguments of the current article.

The asymptotics for growing d in (1.7) is the same as in the corresponding result for the usual Anderson model; moreover, a heuristic analysis of resonances suggests that $d \log d$ is the true order of growth of the threshold (cf. Abou-Chacra *et al.* [1, (6.17)–(6.18)] and a recent rigorous result of Bapst [9] pertaining to the Anderson model on a tree). Also, the arguments of [37] can be applied in the current context, to express the constant C in terms of the connectivity constant of self-avoiding walk on $\mathbb{Z}^d$.

Wegner estimates.

Next we discuss Wegner estimates for a class of models, which contains, in particular, the models (1.1) and certain Gaussian band matrices. For a given Hermitian matrix H and an interval $I \subset \mathbb{R}$ denote

$$\mathcal{N}(H, I) = \#\{\text{eigenvalues of } H \text{ in } I\}\,.$$

Also denote by $|I|$ the length of I.

Estimates on the density of states (cf. (1.11) and the subsequent remark below) were first obtained, in the context of Schrödinger operators, by Wegner [43], who proved the following:

Let H_0 be a $k \times k$ Hermitian matrix, and let $H = H_0 + V$, where V is a random diagonal matrix with entries independently sampled from a bounded probability density p on $\mathbb{R}$. Then

$$\mathbb{E}\mathcal{N}(H, I) \leq \|p\|_\infty \, k \, |I|\,, \quad I \text{ is an interval in } \mathbb{R}\,. \tag{1.8}$$

The original motivation of Wegner was to rule out the divergence of the density of states at the mobility edge. Since then, estimates on the mean number of eigenvalues in an interval, commonly referred to as Wegner estimates, have found numerous applications in the mathematical study of random operators (where they allow to handle resonances).

The original estimate (1.8) can be applied to the finite-volume restrictions of the models (1.2), where it provides the sharp dependence on the volume and on the size of the interval, but not on the number of orbitals. We prove a form of (1.8) tailored to the models (1.2), with the sharp dependence on N. We formulate the result in a more general form, which applies also to Gaussian band matrices.

For positive integers $k, N_1, \ldots, N_k$, we consider a random square matrix of dimension $\sum_{j=1}^{k} N_j$ which has the form

$$H = H_0 + \bigoplus_{j=1}^{k} V(j) = H_0 + \left(\begin{array}{c|c|c|c|c} V(1) & 0 & 0 & \cdots & 0 \\ \hline 0 & V(2) & 0 & \cdots & 0 \\ \hline 0 & 0 & V(3) & \cdots & 0 \\ \hline & & & & \\ \hline 0 & 0 & 0 & \cdots & V(k) \end{array}\right) \tag{1.9}$$

in which H_0 is deterministic and the matrices $(V(j))$ are random and independent, with $V(j)$ of size $N_j \times N_j$. We assume that either the distribution of each $V(j)$ is given by the GOE, and that H_0 is real symmetric (orthogonal case), or that the distribution of each $V(j)$ is given by the GUE and H_0 is Hermitian (unitary case). We refer to matrices thus defined as deformed block-Gaussian matrices.

Theorem 2. There exists a constant $C > 0$ such that the following holds. Let H be a deformed block-Gaussian matrix as in (1.9), in either the orthogonal case or the unitary case. Then, for any interval $I \subset \mathbb{R}$,

$$\mathbb{E}\mathcal{N}(H, I) \le C \sum_{j=1}^{k} N_j \, |I| \, . \tag{1.10}$$

□

The unitary case of Theorem 2 was also recently proved by Pchelin [34]. Pchelin relies on a single-block estimate ($k = 1$, cf. Proposition 2.1 and (2.2) below) which he proves in the unitary case. Possibly, his argument can rely additionally on the orthogonal case of the single-block estimate, as proved in [4], and yield an alternative proof of Theorem 2 in full generality.

Our proofs of Theorem 2 and Theorem 3 below rely on a representation formula for $\mathcal{N}(H, I)$ in terms of similar quantities for single blocks. This formula, presented in Proposition 2.4 (see also Remark 2.5), may possibly be of use elsewhere.

One can obtain complementary bounds to Theorem 2; see Section 4.

Application # 1: orbital model. Going back to the orbital operators (1.1), the theorem applies to the restriction of each of them to a finite volume. For integers $L \geq 0$ and $d \geq 1$ we write

$$\Lambda_L^d := \{-L, -L+1, \ldots, L\}^d.$$

Corollary 1.3. There exists a constant $C > 0$ such that the following holds. Let H be as in (1.1) in either the orthogonal case or the unitary case, and let H_Λ be the restriction of H to a finite volume $\Lambda \subset \mathbb{Z}^d$ as in (1.4). Then

$$\mathbb{E}\mathcal{N}(H_\Lambda, I) \leq CN|\Lambda||I|$$

for any interval $I \subset \mathbb{R}$. In particular, if the limiting measure

$$\rho_H(\cdot) = \lim_{L\to\infty} \frac{\mathbb{E}\mathcal{N}(H_{\Lambda_L^d}, \cdot)}{N(2L+1)^d} \tag{1.11}$$

exists, then it has a density (called the density of states of H) which is bounded uniformly in N. □

We remark that according to general results pertaining to metrically transitive [= ergodic] operators, see Pastur and Figotin [32] or Aizenman and Warzel [6], the limiting measure in (1.11) exists for the models (1.2) and, more generally, whenever the distribution of the hopping terms $W(x, y)$ depends only on $x - y$.

Proof of Corollary 1.3. Condition on the hopping terms $W(x, y)$ and apply Theorem 2 with $k = |\Lambda|$ and all N_j equal to N. ■

Let us briefly discuss related previous results. Constantinescu *et al.* [15] derived an integral representation of the density of states for a class of locally gauge-invariant operators including H^{Weg}. Using this representation, they proved, for a specific model slightly outside the class (1.1), that the density of states is analytic, uniformly in N, in a certain range of parameters. In the case of $d = 1$, further results pertaining to the density of states were obtained by Constantinescu [14] using supersymmetric transfer matrices.

The integrated density of states (the cumulative distribution function of the measure ρ_H from (1.11)) was studied by Khorunzhiy and Pastur [27], who established, for a wide class of orbital models, an asymptotic expansion in inverse powers of N; see further Pastur [31] and the book [33, Section 17.3] of Pastur and Shcherbina.

Application # 2: Gaussian band matrices. We proceed to define a class of Gaussian random matrices to which the results can be applied, and which contains the class of Gaussian band matrices. We say that a random variable is complex Gaussian if its real and imaginary parts are independent real Gaussian random variables with equal variance.

Definition 1.4. Let $L \geq 0, d \geq 1$ be integers and let $\psi : \mathbb{Z}^d \to [0,\infty)$ satisfy $\psi(-r) = \psi(r)$. A Gaussian random matrix H_L with domain Λ_L^d and shape function ψ is an Hermitian $(2L+1)^d \times (2L+1)^d$ random matrix, whose rows and columns are indexed by the elements of Λ_L^d, having the form

$$H_L = \frac{X_L + X_L^*}{\sqrt{2}},$$

where the entries of the matrix X_L are either independent real Gaussian (orthogonal case) or independent complex Gaussian (unitary case), having zero mean and satisfying

$$\mathbb{E}|X_L(x,y)|^2 = \psi(x-y), \quad x, y \in \Lambda_L^d. \qquad \square$$

We remark that an equivalent way to specify the covariance structure of H_L in the above definition is via the formula

$$\mathbb{E} H_L(x,y)\overline{H_L(x',y')} = \psi(x-y) \times \begin{cases} \mathbb{1}_{x=x',y=y'} + \mathbb{1}_{x=y',y=x'} \,, & \text{orthogonal case} \\ \mathbb{1}_{x=x',y=y'} \,, & \text{unitary case}\,. \end{cases} \tag{1.12}$$

Remark 1.5. We note for later use that, in our normalisation, an $N \times N$ random GOE (GUE) matrix has the same distribution as $\frac{X+X^*}{\sqrt{2N}}$ where the entries of the matrix X are independent real (complex) Gaussian with zero mean and with $\mathbb{E}|X(x,y)|^2 = 1$ for all x, y. $\square$

Theorem 2 implies a Wegner estimate for the Gaussian random matrices thus defined. We write $\|v\|_\infty$ for the ℓ^∞ norm of a vector v.

Corollary 1.6. There exists a constant $C > 0$ such that the following holds. Let H_L be a Gaussian random matrix with domain Λ_L^d and shape function ψ in either the orthogonal case or the unitary case. Suppose that

$$\psi(r) \geq \frac{1}{(W+1)^d} \quad \text{when} \quad \|r\|_\infty \leq 2\min(W,L) \tag{1.13}$$

for some integer W satisfying $0 \leq W \leq 2L$. Then for any interval $I \subset \mathbb{R}$,

$$\mathbb{E}\mathcal{N}(H_L, I) \leq C(2L+1)^d |I| \;. \tag{1.14}$$

In particular, assuming (1.13) holds for some integer $W \geq 0$, the measure

$$\rho(\cdot) = \lim_{L\to\infty} \frac{\mathbb{E}\mathcal{N}(H_L,\cdot)}{(2L+1)^d}$$

has a density, the density of states, which is uniformly bounded by C. □

The corollary is particularly interesting in the case when W is a large parameter, $L \gg W$, and $\psi(r)$ is small for $\|r\|_\infty \gg W$; in this case H_L is informally called a Gaussian band matrix of bandwidth W. One way to construct such matrices is to choose, slightly modifying the definition used by Erdős and Knowles [22], the shape function ψ of the form

$$\psi(r) = \frac{\phi(\frac{r}{W})}{W^d}, \quad r \in \mathbb{Z}^d \tag{1.15}$$

for an almost everywhere continuous function $\phi : \mathbb{R}^d \to [0,\infty)$ satisfying $\phi(-r) = \phi(r)$ and $0 < \int \phi(r)\mathrm{d}r < \infty$. If H_L is constructed in this way, and if

$$\phi(\rho) \geq \delta \quad \text{for} \quad \|\rho\|_\infty \leq \epsilon \tag{1.16}$$

with some $0 < \epsilon \leq \frac{4L}{W}$ and $\delta > 0$, then, for any interval $I \subset \mathbb{R}$,

$$\mathbb{E}\mathcal{N}(H_L, I) \leq K\,(2L+1)^d |I| \;, \tag{1.17}$$

where $K = C\sqrt{\frac{1}{\delta}\left(\frac{2}{\epsilon}\right)^d}$. This follows from Corollary 1.6 applied to the matrix $\sqrt{\frac{1}{\delta}\left(\frac{2}{\epsilon}\right)^d} H_L$ with $\lfloor \frac{\epsilon W}{2} \rfloor$ in place of W.

Another example of Gaussian band matrices, in which Corollary 1.6 can be applied to deduce (1.17) with a constant K independent of W, is given by

$$\begin{aligned} &\mathbb{E} H_L(x,y)\overline{H_L(x',y')} \\ &= (-W^2\Delta + \mathbb{1})^{-1}(x,y) \times \begin{cases} \mathbb{1}_{x=x',y=y'} + \mathbb{1}_{x=y',y=x'} \;, & \text{orthogonal case} \\ \mathbb{1}_{x=x',y=y'} \;, & \text{unitary case} \;, \end{cases} \end{aligned} \tag{1.18}$$

where Δ is the discrete Laplacian on $\mathbb{Z}^d$, $d \geq 1$.

This example was studied by Disertori, Pinson, and Spencer [18], who proved an estimate of the form (1.17) for the unitary case in dimension $d = 3$. Very recently, a parallel result for $d = 1$ was proved by M. and T. Shcherbina [38], and for $d = 2$ — by Disertori and Lager [17].

To the best of our knowledge, these are the only previously known estimates of the form (1.17) for any kind of band matrices which are valid for arbitrarily short intervals I uniformly in W; see further [40, Section 3] for a discussion of the problem. We remark that the methods of [17, 18] and [38] allow to go beyond a uniform bound on the density of states, and provide a differentiable asymptotic expansion for it in powers of W^{-2}. On the other hand, these methods make essential use of the particular structure (1.18).

In a generality similar to that of Definition 1.4, Bogachev *et al.* [10] and Khorunzhiy *et al.* [26] found the limit of $\mathcal{N}(H_L, I)/(2L+1)^d$ (with or without the expectation) for a fixed interval I as $W, L \to \infty$; this limit is bounded by a constant times the length of I. The results of Erdős *et al.* [23] (and, in a slightly different setting, of [39]) yield an estimate of the form (1.17) for intervals I of length $|I| \geq W^{-1+\epsilon}$.

Proof of Corollary 1.6. Using the assumption that $0 \leq W \leq 2L$ we may partition $\{-L, -L+1, \ldots, L\}$ into disjoint discrete intervals I_j, $1 \leq j \leq \ell$, satisfying $W + 1 \leq |I_j| \leq 2W + 1$ for all j (if $W \geq L$ then the partition necessarily has $\ell = 1$ and $|I_1| = 2L + 1$). Correspondingly, write

$$\Lambda_L^d = \biguplus_{j=1}^{\ell^d} B_j \tag{1.19}$$

where the B_j are all Cartesian products of the form $J_1 \times J_2 \times \cdots \times J_d$ where each J_i is one of the intervals I_j.

Now consider the matrix H_L as a block matrix, where the partition of the index set Λ_L^d into blocks is given by (1.19). The assumption (1.13), the fact that the entries of H_L are Gaussian and the observation in Remark 1.5 allow us to write

$$H_L = H_L^0 + V_L \tag{1.20}$$

where V_L is a block-diagonal matrix, with the diagonal blocks distributed as GOE in the orthogonal case and as GUE in the unitary case, and where H_L^0 is an Hermitian matrix, independent of V_L, with jointly Gaussian entries which are real in the orthogonal case.

Thus, conditioning on H_L^0, the estimate (1.14) follows from Theorem 2 applied with $k = \ell^d$ and $N_j = |B_j|$. ■

Minami estimates.

In the same setting as (1.8), Minami established [29] the bound:

$$\mathbb{E}\mathcal{N}(H,I)(\mathcal{N}(H,I)-1) \leq (C\|p\|_\infty|\Lambda|\,|I|)^2\,. \tag{1.21}$$

The bound (1.21) rules out attraction between eigenvalues in the local regime; it is a key step in Minami's proof of Poisson statistics for the Anderson model in the regime of Anderson localisation. Subsequently, additional proofs and generalisations of (1.21) were found, among which we mention the argument of Combes *et al.* [13].

The next result is a counterpart of (1.21) in our block setting. As in Theorem 2, the central feature is the dependence on the sizes of the blocks.

Theorem 3. There exists $C > 0$ such that the following holds. Let H be a deformed block-Gaussian matrix as in (1.9), in either the orthogonal case or the unitary case. Then, for any integer $m \geq 1$ and interval $I \subset \mathbb{R}$,

$$\mathbb{E}\prod_{\ell=0}^{m-1}(\mathcal{N}(H,I)-\ell) \leq \left(C\sum_{j=1}^{k} N_j\,|I|\right)^m\,, \tag{1.22}$$

and, consequently,

$$\mathbb{P}\{\mathcal{N}(H,I)\geq m\} \leq \frac{1}{m!}\left(C\sum_{j=1}^{k} N_j\,|I|\right)^m\,. \qquad \square$$

The case $m = 1$ in the theorem is the Wegner estimate discussed in Theorem 2 whereas the cases $m \geq 2$ are Minami-type estimates.

Localisation for one-dimensional band matrices.

Band matrices in one dimension ($d = 1$) have been studied extensively in the physics literature as a simple model in which the quantum dynamics exhibits crossover from quantum diffusion to localisation, see [11, 12, 25]. Based on those works, the following crossover is expected: considering band matrices of dimension L and bandwidth W, when $W \ll \sqrt{L}$, each eigenvector has appreciable overlap with a vanishingly small fraction of the standard basis vectors in the large L limit, whereas for $W \gg \sqrt{L}$ a typical eigenvector has overlap of order $1/\sqrt{L}$ with most standard basis vectors. A related conjecture states that the i,j-entry of the resolvent should decay as $\exp(-C|i-j|W^{-2})$ for $W \ll \sqrt{L}$.

In [36], one of us studied the localisation side of this problem. In that paper it was shown that certain ensembles of random matrices whose entries vanish outside a band of width W around the diagonal satisfy a localisation condition in the limit that the size of the matrix L tends to infinity such that $W^8/L \to 0$. For Gaussian band matrices, our present work settles [36, Problem 2] in the positive, thereby allowing to improve the result there slightly by replacing the exponent 8 with the exponent 7 (which is still a bit away from the expected optimal exponent 2).

Theorem 4. Let H_L be a Gaussian random matrix with domain $\Lambda_L = \{-L, -L+1, \ldots, L\}$ and shape function ψ as in Definition 1.4 in either the orthogonal case or the unitary case. Let W be an integer dividing $2L+1$ and suppose that ψ is the sharp cutoff function

$$\psi(r) = \begin{cases} \frac{1}{W} & |x| < W \\ 0 & |x| \geq W \end{cases} . \tag{1.23}$$

Then, given $\rho > 0$ and $s \in (0,1)$ there are $A < \infty$ and $\alpha > 0$ such that

$$\mathbb{E}\left(\left|(H_L - \lambda)^{-1}(i,j)\right|^s\right) \ \leq \ A W^{\frac{s}{2}} \mathrm{e}^{-\alpha \frac{|i-j|}{W^7}} \tag{1.24}$$

for all $\lambda \in [-\rho, \rho]$ and all $i,j \in \{-L, \ldots, L\}$. □

Remark 1.7. The theorem implies (using the resolvent identity) that a similar estimate holds without the assumption $2L+1 \equiv 0 \mod W$. □

2 Proof of the Theorems

In this section, we prove the main results of the article. We use the following result from [4], where the object of study was the regularizing effect of adding a Gaussian random matrix to a given deterministic matrix. The GUE case of (2.2) was also proved by Pchelin [34].

Proposition 2.1. [4, Theorem 1 and Remark 2.2]

If either: A is an $N \times N$ real symmetric matrix, $v \in \mathbb{R}^N$, and V is sampled from GOE,
or: A is an $N \times N$ Hermitian matrix, $v \in \mathbb{C}^N$, and V is sampled from GUE,

then the matrix $A + V$ satisfies the bounds:

$$\mathbb{P}\left\{\|(A+V)^{-1}v\| \geq t\sqrt{N}\|v\|\right\} \leq \frac{C}{t}, \quad t \geq 1, \tag{2.1}$$

$$\mathbb{E}\mathcal{N}(A+V, I) \leq CN|I|, \quad I \text{ is an interval in } \mathbb{R} \tag{2.2}$$

with a constant $C < \infty$ which is uniform in N, A, and v. Moreover, the following stronger version of (2.1) holds: almost surely,

$$\mathbb{P}\left\{\|(A+V)^{-1}v\| \geq t\sqrt{N}\|v\| \;\middle|\; \hat{P}_{v^\perp} V \hat{P}^*_{v^\perp}\right\} \leq \frac{C}{t}, \quad t \geq 1, \tag{2.3}$$

where $\hat{P}_{v^\perp} : \mathbb{C}^N \to \mathbb{C}^N/\mathbb{C}v \simeq v^\perp$ is the canonical projection, and $v^\perp$ is the orthogonal complement to v. □

In the setting of Proposition 2.1 we note the following consequence of (2.1),

$$\mathbb{E}\|(A+V)^{-1}v\|^s \leq \frac{C_0 N^{\frac{s}{2}}}{1-s}\|v\|^s, \quad 0 < s < 1, \tag{2.4}$$

where C_0 is an absolute constant (uniform in all the parameters of the problem).

Remark 2.2. Our proofs of Theorem 1 (localisation), Theorem 2 (Wegner-type estimate) and Theorem 3 (Minami-type estimate) rely on the Gaussian structure of the underlying random matrix ensembles only through Proposition 2.1 (and the simple observation of Remark 2.3 below). Thus, if an extension of the proposition to other random matrix ensembles is found, corresponding extensions of our theorems will follow. □

Remark 2.3. For the random matrix models discussed in our theorems, any given $\lambda \in \mathbb{R}$ is almost surely not an eigenvalue. This follows, for instance, from the following observation: if H is a random matrix satisfying that for any $\mu \in \mathbb{R}$, the distribution of H and the distribution of $H - \mu$ are mutually absolutely continuous then any given $\lambda \in \mathbb{R}$ is almost surely not an eigenvalue of H. □

Proof of Theorem 1. Denote by $G_\lambda[\tilde{H}] = (\tilde{H} - \lambda)^{-1}$ the resolvent of an operator $\tilde{H}$. For $\tilde{x} \in \tilde{\Lambda} \subset \Lambda$, let $\mathcal{F}_{\tilde{\Lambda},\tilde{x}}$ be the σ-algebra generated by all $H_{\tilde{\Lambda}}(w, w')$, where $w, w' \in \tilde{\Lambda}$ and $(w, w') \neq (\tilde{x}, \tilde{x})$.

Observe the following corollary of the Schur–Banachiewicz formula for block matrix inversion: for any $\tilde{x} \in \tilde{\Lambda} \subset \Lambda$,

$$G_\lambda[H_{\tilde{\Lambda}}](\tilde{x}, \tilde{x}) = (V(\tilde{x}) - \lambda - \Sigma)^{-1} , \tag{2.5}$$

where Σ is measurable with respect to $\mathcal{F}_{\tilde{\Lambda},\tilde{x}}$. Consequently, by (2.4), almost surely,

$$\mathbb{E}\left[\left\| G_\lambda[H_{\tilde{\Lambda}}](\tilde{x}, \tilde{x})\tilde{v} \right\|^s \mid \mathcal{F}_{\tilde{\Lambda},\tilde{x}} \right] \le \frac{C_0 N^{\frac{s}{2}}}{1-s} \mathbb{E}\|\tilde{v}\|^s \tag{2.6}$$

whenever $\tilde{v}$ is a random vector which is measurable with respect to $\mathcal{F}_{\tilde{\Lambda},\tilde{x}}$.

Next, we use the following representation of $G_\lambda[H_\Lambda](x, y)$:

$$\begin{aligned} G_\lambda[H_\Lambda](x, y) = \sum_{k \ge \|x-y\|_1} (-1)^k \sum_{\pi \in \Pi_k(x,y)} G_\lambda[H_\Lambda](\pi_0, \pi_0) W(\pi_0, \pi_1) G_\lambda[H_{\Lambda \setminus \{\pi_0\}}](\pi_1, \pi_1) \\ W(\pi_1, \pi_2) G_\lambda[H_{\Lambda \setminus \{\pi_0,\pi_1\}}](\pi_2, \pi_2) \cdots W(\pi_{k-1}, \pi_k) G_\lambda[H_{\Lambda \setminus \{\pi_0,\pi_1,\cdots,\pi_{k-1}\}}](\pi_k, \pi_k) , \end{aligned} \tag{2.7}$$

where for $y = x$ the right-hand side is interpreted as $G_\lambda[H_\Lambda](x, x)$, and for $y \ne x$ the collection $\Pi_k(x, y)$ includes all tuples of pairwise distinct vertices $\pi_0, \pi_1, \cdots, \pi_k \in \Lambda$ such that $x = \pi_0 \sim \pi_1 \sim \pi_2 \cdots \sim \pi_k = y$.

Indeed, the representation is tautological for $y = x$, and for $y \ne x$ it follows by iterating the equality

$$G_\lambda[H_\Lambda](x, y) = -\sum_{\pi_1 \sim x} G_\lambda[H_\Lambda](x, x) W(x, \pi_1) G_\lambda[H_{\Lambda \setminus \{x\}}](\pi_1, y) .$$

The latter is in turn a corollary of the second resolvent identity applied to the operators $H_\Lambda - \lambda$ and $H_\Lambda^x - \lambda$, where H_Λ^x is obtained from H_Λ by setting the blocks $W(x, x')$ and $W(x', x)$ to 0 for all $x' \sim x$.

Now we turn to the proof of the theorem. We derive from (2.7) using the triangle inequality and $|a + b|^s \le |a|^s + |b|^s$ that

$$\begin{aligned} \|G_\lambda[H_\Lambda](x, y)v\|^s \le \sum_{k \ge \|x-y\|_1} \sum_{\pi \in \Pi_k(x,y)} \Big\| G_\lambda[H_\Lambda](\pi_0, \pi_0) W(\pi_0, \pi_1) G_\lambda[H_{\Lambda \setminus \{\pi_0\}}](\pi_1, \pi_1) \\ W(\pi_1, \pi_2) G_\lambda[H_{\Lambda \setminus \{\pi_0,\pi_1\}}](\pi_2, \pi_2) \cdots W(\pi_{k-1}, \pi_k) G_\lambda[H_{\Lambda \setminus \{\pi_0,\pi_1,\cdots,\pi_{k-1}\}}](\pi_k, \pi_k) v \Big\|^s . \end{aligned} \tag{2.8}$$

To bound the expectation of a term in (2.8), we repeatedly use (2.6) and the inequality

$$\mathbb{E}\|W(\tilde{x}, \tilde{x}')\|_{\mathrm{op}}^s \le g_{\mathrm{eff}}^s \tag{2.9}$$

from (1.5). We obtain for the term in (2.8) corresponding to a single $\pi \in \Pi_k(x,y)$:

$$\mathbb{E}\left\| G_\lambda[H_\Lambda](\pi_0,\pi_0)W(\pi_0,\pi_1)\cdots G_\lambda[H_{\Lambda\setminus\{\pi_0,\pi_1,\cdots,\pi_{k-1}\}}](\pi_k,\pi_k)v\right\|^s \le \left(\frac{C_0N^{\frac{s}{2}}}{1-s}\right)^{k+1} g_{\mathrm{eff}}^{sk}\ \|v\|^s .$$

The cardinality of $\Pi_k(x,y)$ does not exceed $(2d)^k$. Therefore

$$\begin{aligned}\mathbb{E}\|G_\lambda[H_\Lambda](x,y)v\|^s &\le \sum_{k\ge\|x-y\|_1}\left(\frac{C_0N^{\frac{s}{2}}}{1-s}\right)^{k+1}\left(2dg_{\mathrm{eff}}^s\right)^k\ \|v\|^s\\ &\le 2\left(\frac{C_0N^{\frac{s}{2}}}{1-s}\right)^{\|x-y\|_1+1}\left(2dg_{\mathrm{eff}}^s\right)^{\|x-y\|_1}\ \|v\|^s\end{aligned}$$

whenever

$$\frac{4C_0dg_{\mathrm{eff}}^sN^{\frac{s}{2}}}{1-s}\le 1 .$$

This is what is claimed in the statement of the theorem, for $C=4C_0$. ■

The proofs of Theorems 2 and 3 are preceded by the following proposition, the purpose of which is to write $\mathcal{N}(H,I)$ as a linear expression involving terms of the form $\mathcal{N}(V+A,J)$, where V is a random matrix sampled from the GOE (or GUE), A is a symmetric (or Hermitian) matrix independent of V and J is an interval in $\mathbb{R}$.

Proposition 2.4. Let H have the form $H=H_0+\oplus_{j=1}^k V(j)$ (as in (1.9)) in which H_0 and all $V(j)$ are (deterministic) Hermitian matrices. Then for any interval I the endpoints of which are not eigenvalues of H,

$$\mathcal{N}(H,I)=\lim_{\eta\to+0}\sum_{j=1}^k\int_I\frac{\mathrm{d}\lambda}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{\pi(1+t^2)}\int_0^{\infty}\frac{2\eta\xi\ d\xi}{(\xi^2+\eta^2)^2}\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),(-\xi,\xi)\Big),$$

where $A(j,\lambda,\eta,t)$ is an Hermitian matrix determined by H_0 and $(V(\ell))_{\ell\ne j}$ (i.e., every matrix element of A is a Borel-measurable function of these variables and λ, η and t). In addition, if H_0 and all $V(j)$ are real, then the matrices A are real as well. □

The exact definition of the matrices $A(j,\lambda,\eta,t)$ is given in (3.1) and (3.2).

Remark 2.5. Each of the integrals

$$\int_I\frac{\mathrm{d}\lambda}{|I|}\ ,\quad \int_{-\infty}^{\infty}\frac{\mathrm{d}t}{\pi(1+t^2)}\ ,\quad \int_0^{\infty}\frac{4\eta\xi^2\,\mathrm{d}\xi}{\pi(\xi^2+\eta^2)^2}$$

equals one. This leads us to introduce the following notation:

$$\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\,\Phi(\lambda,t,\xi;\eta) = \int_I \frac{\mathrm{d}\lambda}{|I|} \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \int_0^{\infty} \frac{4\eta\xi^2\,\mathrm{d}\xi}{\pi(\xi^2+\eta^2)^2}\,\Phi(\lambda,t,\xi;\eta)\;.$$

With this notation and assuming $0 < |I| < \infty$, the conclusion of Proposition 2.4 takes the form:

$$\frac{1}{|I|}\mathcal{N}(H,I) = \lim_{\eta\to+0}\sum_{j=1}^{k}\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\;\frac{1}{2\xi}\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),(-\xi,\xi)\Big). \tag{2.10}$$

□

We prove Theorems 2 and 3 using Proposition 2.4, and defer the proof of the proposition to the next section.

Proof of Theorem 2. In the setting of the theorem, the end points of any fixed interval I are almost surely not eigenvalues of H (see Remark 2.3). Therefore Proposition 2.4 is applicable almost surely, and (2.10) yields:

$$\begin{aligned}\frac{1}{|I|}\mathbb{E}\mathcal{N}(H,I) &= \mathbb{E}\lim_{\eta\to+0}\sum_{j=1}^{k}\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\;\frac{1}{2\xi}\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),(-\xi,\xi)\Big)\\ &\leq \lim_{\eta\to+0}\sum_{j=1}^{k}\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\;\frac{1}{2\xi}\mathbb{E}\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),(-\xi,\xi)\Big)\\ &\leq \lim_{\eta\to+0}\sum_{j=1}^{k}\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\;CN_j = C\sum_{j=1}^{k}N_j\;.\end{aligned}$$

The first inequality follows from the Fatou lemma and the second one is an application of the single-block bound (2.2) with $|I| = 2\xi$. ■

The proof of Theorem 3 also uses formula (2.10) as a starting point, and proceeds following arguments similar to those used in the proof of [4, Theorem 2] (which is the $k = 1$ case of Theorem 3); these arguments are, in turn, inspired by the work of Combes *et al.* [13]. We start with the following simple lemma (see e.g., [4, Lemma 3.1 and (3.6)] for a slightly stronger version featuring the Frobenius norm in place of the operator norm).

Lemma 2.6. Let A be an $N\times N$ deterministic Hermitian matrix.

If either: v is uniformly distributed on the sphere $\mathbb{S}^{N-1}_{\mathbb{R}} = \{w\in\mathbb{R}^N : \|w\| = 1\}$
and A is real,

or: v is uniformly distributed on the sphere $\mathbb{S}^{N-1}_{\mathbb{C}} = \{w\in\mathbb{C}^N : \|w\| = 1\}$, then

$$\mathbb{P}\left\{\|Av\| \le \frac{\epsilon}{\sqrt{N}}\|A\|_{\mathrm{op}}\right\} \le 5\epsilon, \quad \epsilon > 0. \qquad \square$$

Consequently, in the same setting, for any non-negative random variable X which is independent of v,

$$\mathbb{E}X \le 2\mathbb{E}\left[X \cdot \mathbb{1}\left\{\|A^{-1}v\| \ge \frac{\|A^{-1}\|_{\mathrm{op}}}{10\sqrt{N}}\right\}\right], \tag{2.11}$$

where $\mathbb{1}\{\Omega\}$ is the indicator of an event Ω.

Proof of Theorem 3. It suffices to prove the theorem for intervals I of length $0 < |I| < \infty$, therefore we tacitly impose this assumption on all intervals which appear in this proof. The argument is by induction on m. Let $C_1 = 10\,C$, where C is the greater among the constants in Theorem 2 and Proposition 2.1. Fix an interval I and the numbers N_j; let $m \ge 2$, and assume, as the induction hypothesis, that

$$\mathbb{E}\prod_{\ell=0}^{m-2}(\mathcal{N}(H,I) - \ell) \le \left(C_1\sum_{j=1}^{k} N_j\,|I|\right)^{m-1}, \tag{2.12}$$

for any deformed block-Gaussian random matrix H of the form (1.9) in either the orthogonal case or the unitary case. Note that the induction base, (2.12) with $m = 2$, follows from Theorem 2.

Let H be a random matrix of the form (1.9) in either the orthogonal case or the unitary case. The formula (2.10) applied to $\mathcal{N}(H,I)$ shows that

$$\prod_{\ell=0}^{m-1}(\mathcal{N}(H,I) - \ell) = |I|\lim_{\eta\to+0}\sum_{j=1}^{k}\mathrm{Ave}^{\eta}_{\lambda,t,\xi}\ \frac{1}{2\xi}\mathcal{N}\Big(V(j) + A(j,\lambda,\eta,t), (-\xi,\xi)\Big)\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I) - \ell)\,.$$

Thus, by the Fatou lemma, it suffices to prove that for any $1 \le j \le k, \lambda, t \in \mathbb{R}$ and $\xi, \eta > 0$,

$$\mathbb{E}\mathcal{N}\Big(V(j) + A(j,\lambda,\eta,t), (-\xi,\xi)\Big)\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I) - \ell)_+ \le 2\xi \cdot C_1 N_j\left(C_1\sum_{i=1}^{k} N_i\,|I|\right)^{m-1}. \tag{2.13}$$

The eigenvalues of $V(j) + A(j,\lambda,\eta,t)$ are simple almost surely, since the distribution of $V(j) + A(j,\lambda,\eta,t)$ is absolutely continuous with respect to the Lebesgue measure on real symmetric matrices (orthogonal case) or Hermitian matrices (unitary case). For each natural M, construct a partition $\{I_M^n\}_{n=1}^{2^M}$ of $(-\xi,\xi)$ into 2^M intervals of equal length. Then, almost surely,

$$\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),(-\xi,\xi)\Big)=\lim_{M\to\infty}\sum_{n=1}^{2^M}\mathbb{1}\left\{\mathcal{N}\Big(V(j)+A(j,\lambda,\eta,t),\,I_M^n\Big)\geq 1\right\}$$
$$=\lim_{M\to\infty}\sum_{n=1}^{2^M}\mathbb{1}\left\{\|(V(j)+A(j,\lambda,\eta,t)-\mathcal{M}(I_M^n))^{-1}\|_{\mathrm{op}}\geq\frac{2}{|I_M^n|}\right\},$$

where we denoted by $\mathcal{M}(J)$ the mid-point of an interval $J\subset\mathbb{R}$. This relation, combined with the monotone convergence theorem (as the partitions are refining when M increases), reduces the desired (2.13) to the following claim: for any interval J,

$$\mathbb{E}\,\mathbb{1}\left\{\|B_{j,J}^{-1}\|_{\mathrm{op}}\geq\frac{2}{|J|}\right\}\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I)-\ell)_+\leq|J|\cdot C_1N_j\left(C_1\sum_{i=1}^{k}N_i\,|I|\right)^{m-1},\qquad(2.14)$$

where we denoted

$$B_{j,J}:=V(j)+A(j,\lambda,\eta,t)-\mathcal{M}(J)\,.$$

Now let v be a random vector, independent of H, which is uniformly distributed on the sphere $\mathbb{S}_{\mathbb{R}}^{N_j-1}$ in the orthogonal case or uniformly distributed on the complex sphere $\mathbb{S}_{\mathbb{C}}^{N_j-1}$ in the unitary case. By first conditioning on H, inequality (2.11) may be applied to show that

$$\begin{aligned}&\mathbb{E}\,\mathbb{1}\left\{\|B_{j,J}^{-1}\|_{\mathrm{op}}\geq\frac{2}{|J|}\right\}\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I)-\ell)_+\\&\leq 2\mathbb{E}\left[\mathbb{1}\left\{\|B_{j,J}^{-1}\|_{\mathrm{op}}\geq\frac{2}{|J|}\right\}\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I)-\ell)_+\cdot\mathbb{1}\left\{\|B_{j,J}^{-1}v\|\geq\frac{\|B_{j,J}^{-1}\|_{\mathrm{op}}}{10\sqrt{N_j}}\right\}\right]\\&\leq 2\mathbb{E}\left[\mathbb{1}\left\{\|B_{j,J}^{-1}v\|\geq\frac{1}{5\sqrt{N_j}|J|}\right\}\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I)-\ell)_+\right].\end{aligned}\qquad(2.15)$$

Denote by $P_j:\mathbb{R}^{\sum_i N_i}\to\mathbb{R}^{N_j}$ the coordinate projection to the space corresponding to $V(j)$, and, for $\tau\in\mathbb{R}$, define the rank-one perturbation

$$H_{v,\tau}=H+\tau P_j^*vv^*P_j\,.$$

The eigenvalues of H and $H_{v,\tau}$ interlace, therefore

$$\prod_{\ell=1}^{m-1}(\mathcal{N}(H,I)-\ell)_+ \le \mathfrak{P} := \lim_{\tau\to+\infty}\prod_{\ell=0}^{m-2}(\mathcal{N}(H_{v,\tau},I)-\ell) \tag{2.16}$$

(the inequality actually holds for any fixed τ). In view of (2.15), our goal (2.14) is reduced to the inequality

$$2\mathbb{E}\left[\mathbb{1}\left\{\|B_{j,J}^{-1}v\| \ge \frac{1}{5\sqrt{N_j}|J|}\right\}\mathfrak{P}\right] \le |J|\cdot C_1 N_j\left(C_1\sum_{i=1}^{k}N_i\,|I|\right)^{m-1}, \tag{2.17}$$

which we now prove.

The following simple fact is central to the argument. For an Hermitian matrix K of dimension r and unit vector $u \in \mathbb{C}^r$, define a matrix K_u of dimension $r-1$ by $K_u = \hat{P}_{u^\perp}K_u\hat{P}^*_{u^\perp}$, where $\hat{P}_{u^\perp}: \mathbb{C}^r \to \mathbb{C}^r/\mathbb{C}u$ is the canonical projection (e.g., if u is the first vector of the standard basis, K_u is the submatrix obtained by removing the first row and column of K_u). Then

$$\lim_{\tau\to\infty}\mathcal{N}(K+\tau uu^*,I) = \mathcal{N}(K_u,I)$$

for any interval I whose endpoints are not eigenvalues of K_u. We apply this identity with $K = H$ and $u = P_j^*v$, and deduce that the random variable $\lim_{\tau\to+\infty}\mathcal{N}(H_{v,\tau},I)$ is measurable with respect to $\hat{P}_{(P_j^*v)^\perp}H\hat{P}^*_{(P_j^*v)^\perp}$. Thus, the "moreover" part of Proposition 2.1 can be applied, yielding

$$\begin{aligned}&2\mathbb{E}\left[\mathbb{1}\left\{\|B_{j,J}^{-1}v\| \ge \frac{1}{5\sqrt{N_j}|J|}\right\}\mathfrak{P}\right]\\ &= 2\mathbb{E}\left(\mathfrak{P}\cdot\mathbb{P}\left[\|B_{j,J}^{-1}v\| \ge \frac{1}{5\sqrt{N_j}|J|}\;\middle|\;\hat{P}_{(P_j^*v)^\perp}H\hat{P}^*_{(P_j^*v)^\perp}\right]\right) \le C_1N_j|J|\mathbb{E}\mathfrak{P}\end{aligned} \tag{2.18}$$

with $C_1 = 10C$. Now note that each of the matrices $H_{v,\tau}$, conditioned on v, has the form (1.9) in the orthogonal or unitary case (corresponding to the case of H). Thus we may apply the Fatou lemma and the induction hypothesis (2.12) to conclude that

$$\mathbb{E}\mathfrak{P} \le \lim_{\tau\to\infty}\mathbb{E}\prod_{\ell=0}^{m-2}(\mathcal{N}(H_{v,\tau},I)-\ell) \le \left(C_1\sum_{i=1}^{k}N_i\,|I|\right)^{m-1}. \tag{2.19}$$

The combination of (2.18) with (2.19) concludes the proof of (2.17) and of the theorem. ■

Proof of Theorem 4. We consider in parallel the cases of orthogonal and unitary symmetry. First, the matrix H_L is of the form

$$H_L = \begin{pmatrix} V_1 & T_1 & 0 & \cdots & \cdots & \cdots & 0 \\ T_1^* & V_2 & T_2 & 0 & \cdots & \cdots & 0 \\ 0 & T_2^* & V_3 & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \ddots & T_{k-1} \\ 0 & 0 & \cdots & \cdots & 0 & T_{k-1}^* & V_k \end{pmatrix}$$

with V_j, $j = 1, \ldots, k$, being $W \times W$ matrices drawn from the GOE (GUE) and T_j, $j = 1, \ldots, k-1$, being lower triangular real (complex) Gaussian matrices. The individual matrices are identically distributed within each family and stochastically independent (within each family and between the families). The matrix dimension is $2L + 1 = kW$. Therefore, we are in the setting of [36, Section 3].

For the rest of the proof, we fix an arbitrary $t \in (s, 1)$ (its value only affects the constants in the estimates). According to [36, Theorem 6], there exist $C, \mu > 0$ such that for any $i, j \in \{-L, -L+1, \ldots, L\}$,

$$\mathbb{E}\left(\left|(H_L - \lambda)^{-1}(i,j)\right|^s\right) \leq CM(W,t)^{\frac{s}{t}} e^{-\mu W^{-2\nu-1}|i-j|}, \tag{2.20}$$

where

$$M(W,t) = \max_{-L \leq i,j \leq L} \mathbb{E}\left(\left|(H_L - \lambda)^{-1}(i,j)\right|^t\right),$$

and

$$\nu \geq \max(2, \zeta + \max(a, 1 + \sigma + 2b))$$

with ζ, a, σ and b certain exponents related to the distribution of the blocks of H_L. We refer to [36] for the definition and discussion of the exponents ζ, a and b. As explained in [36, Section 5], for the Gaussian matrices considered here we can take $\zeta = 2$, $a = 0$ and $b = 0$.

The key improvement afforded in the present work comes from the exponent σ, which is related to the Wegner estimate, namely σ is such that for any $R > 1$, real symmetric (Hermitian) $W \times W$ matrices A, B and a real (complex) arbitrary $W \times W$ matrix D,

$$\mathbb{P}\left\{\left\|(V-A)^{-1}\right\| > RW^{1+\sigma}\right\} \le \kappa \frac{1}{R}$$

and

$$\mathbb{P}\left\{\left\|\begin{pmatrix} V-A & D \\ D^* & V'-B \end{pmatrix}^{-1}\right\| > RW^{1+\sigma}\right\} \le 2\kappa \frac{1}{R},$$

where V, V' are independent $W \times W$ random matrices with the GOE (GUE) distribution. Theorem 2, applied with one or two diagonal blocks ($k = 1, 2$), ensures that these estimates hold with $\sigma = 0$ (in the single block case it suffices to use the result of [4] stated here as Proposition 2.1). According to a Wegner-type estimate of [3, Theorem II.1],

$$M(W, t) \le C_t W^{\frac{t}{2}}.$$

Plugging this estimate into (2.20) with $\sigma = 0$, we obtain the claim. ■

3 Proof of Proposition 2.4

We start with a preparatory lemma which is a consequence of the Poisson integral formula.

Lemma 3.1. Let X, Y be Hermitian matrices such that Y is negative semi-definite, and let $\eta > 0$. Then

$$\Im \operatorname{tr}(X + iY - i\eta)^{-1} = \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \Im \operatorname{tr}(X + tY - i\eta)^{-1}. \qquad \square$$

Proof. Consider the function

$$\phi(\xi) = \operatorname{tr}(X + \xi Y - i\eta)^{-1}, \quad \xi \in \mathbb{C}, \quad \Im\xi \ge 0.$$

Observe that, for ξ as above and any non-zero vector $\psi \in \mathbb{C}^N$,

$$\Im\langle (X + \xi Y - i\eta)\psi, \psi\rangle \le -\eta\|\psi\|^2 < 0$$

(where N is the common dimension of X and Y, and $\langle\cdot,\cdot\rangle$ is the scalar product on $\mathbb{C}^N$, linear in the first component and anti-linear in the second one). By an elementary linear-algebreaic argument, ϕ is holomorphic in its domain of definition and, in particular,

$\Im\phi$ is harmonic. Also,

$$\Im\phi(\xi) = \frac{1}{2i}\,\mathrm{tr}\left\{(X + \xi Y - i\eta)^{-1} - (X + \bar{\xi} Y + i\eta)^{-1}\right\}$$

is positive, and

$$\limsup_{y\to+\infty} \Im\phi(iy) \le \sup_{y>0} \Im\phi(iy) \le N\eta^{-1} < \infty\ ,$$

since $\|(X+iyY-i\eta)^{-1}\|_{\mathrm{op}} \le \frac{1}{\eta}$. Therefore (see e.g., [28, Chapter 14]) $\Im\phi$ admits the Poisson representation

$$\Im\phi(i) = \int_{-\infty}^{+\infty} \frac{\Im\phi(t)\mathrm{d}t}{\pi(1+t^2)} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \Im\mathrm{tr}(X + tY - i\eta)^{-1}. \qquad \blacksquare$$

We proceed with the proof of Proposition 2.4. Let H have the form $H = H_0 + \oplus_{j=1}^{k} V(j)$ (as in (1.9)) in which H_0 and all $V(j)$ are (deterministic) Hermitian matrices and suppose that $V(j)$ is of size $N_j \times N_j$. Denote by $P_j : \mathbb{R}^{\sum_i N_i} \to \mathbb{R}^{N_j}$ the coordinate projection to the space corresponding to $V(j)$; also denote by $Q_j : \mathbb{R}^{\sum_i N_i} \to \mathbb{R}^{\sum_{i\neq j} N_i}$ the coordinate projection to the orthogonal subspace to the range of P_j. Let

$$A(j) = P_j H_0 P_j^*\ , \quad B(j) = Q_j H P_j^*\ , \quad C(j) = Q_j H Q_j^* \tag{3.1}$$

(note that $A(j)$ is defined with H_0 rather than H) and define, for $\lambda, t \in \mathbb{R}$ and $\eta > 0$,

$$A(j,\lambda,\eta,t) = -\lambda + A(j) - B(j)^*(C(j) - \lambda + t\eta)((C(j) - \lambda)^2 + \eta^2)^{-1}B(j)\ . \tag{3.2}$$

The Perron–Stieltjes inversion formula [7, Addenda to Chapter III], using our assumption that the endpoints of I are not eigenvalues of H, implies that

$$\mathcal{N}(H,I) = \lim_{\eta\to+0} \int_I \frac{\mathrm{d}\lambda}{\pi} \Im\,\mathrm{tr}(H - \lambda - i\eta)^{-1}\ . \tag{3.3}$$

Now, for any $\lambda \in \mathbb{R}$ and $\eta > 0$ the integrand may be rewritten using the Schur–Banachiewicz inversion formula,

$$\begin{aligned}\mathrm{tr}(H-\lambda-i\eta)^{-1} &= \sum_{j=1}^{k} \mathrm{tr}\, P_j (H-\lambda-i\eta)^{-1} P_j^* \\ &= \sum_{j=1}^{k} \mathrm{tr}\left(V(j) - \lambda - i\eta + A(j) - B(j)^*(C(j)-\lambda-i\eta)^{-1}B(j)\right)^{-1}\ .\end{aligned}$$

This expression, in turn, may be rewritten as follows. Denoting, for each $1 \le j \le k$,

$$Z(j) = -\lambda + A(j) - B(j)^*(C(j) - \lambda - i\eta)^{-1}B(j) \ ,$$

we may define the Hermitian matrices

$$X(j) = \frac{Z(j) + Z(j)^*}{2} = -\lambda + A(j) - B(j)^*(C(j) - \lambda)((C(j) - \lambda)^2 + \eta^2)^{-1}B(j) \ ,$$
$$Y(j) = \frac{Z(j) - Z(j)^*}{2i} = -\eta B(j)^*((C(j) - \lambda)^2 + \eta^2)^{-1}B(j)$$

and conclude that

$$\operatorname{tr}(H - \lambda - i\eta)^{-1} = \sum_{j=1}^{k} \operatorname{tr}(V(j) + X(j) + iY(j) - i\eta)^{-1} \ .$$

The matrix $Y(j)$ is explicitly negative semi-definite, therefore Lemma 3.1 implies that

$$\begin{aligned}\Im \operatorname{tr}(H - \lambda - i\eta)^{-1} &= \sum_{j=1}^{k} \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \Im \operatorname{tr}(V(j) + X(j) + tY(j) - i\eta)^{-1} \\ &= \sum_{j=1}^{k} \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \Im \operatorname{tr}(V(j) + A(j,\lambda,\eta,t) - i\eta)^{-1} \ .\end{aligned}$$

Plugging this equality into (3.3) shows that

$$\mathcal{N}(H, I) = \lim_{\eta \to +0} \sum_{j=1}^{k} \int_I \frac{\mathrm{d}\lambda}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\pi(1+t^2)} \Im \operatorname{tr}(V(j) + A(j,\lambda,\eta,t) - i\eta)^{-1} \ . \tag{3.4}$$

To conclude the proof of the proposition, we use the following form of the spectral theorem:

$$\operatorname{tr} f(V(j) + A(j,\lambda,\eta,t)) = \int_{-\infty}^{\infty} f(\xi) \mathrm{d}\mathcal{N}(V(j) + A(j,\lambda,\eta,t), (-\infty, \xi)) \ ,$$

where the (Stieltjes) integral is with respect to the ξ variable, and $\mathcal{N}(V(j) + A(j,\lambda,\eta,t), (a,b))$ denotes the number of eigenvalues of $V(j) + A(j,\lambda,\eta,t)$ in the interval (a,b). Plugging in the even function $f(\xi) = \Im(\xi - i\eta)^{-1} = \eta/(\xi^2 + \eta^2)$, we obtain that

$$\Im \operatorname{tr}(V(j) + A(j,\lambda,\eta,t) - i\eta)^{-1} = \int_0^{\infty} \frac{\eta}{\xi^2 + \eta^2} \, \mathrm{d}\mathcal{N}(V(j) + A(j,\lambda,\eta,t), (-\xi, \xi)) \ ,$$

Integrating by parts, we have:

$$\int_0^\infty \mathrm{d}\mathcal{N}(V(j) + A(j,\lambda,\eta,t), (-\xi,\xi)) \frac{\eta}{\xi^2+\eta^2} = \int_0^\infty \mathcal{N}\Big(V(j) + A(j,\lambda,\eta,t), (-\xi,\xi)\Big) \frac{2\eta\xi\,\mathrm{d}\xi}{(\xi^2+\eta^2)^2} \,.$$

The last two displayed equations together with (3.4) establish the proposition.

4 Concluding Remarks

Second-order perturbation theory.

One heuristic explanation for the scaling (1.3) is provided by second-order perturbation theory. We sketch the argument for H^{bA}; similar considerations apply to the Wegner N-orbital operator H^{Weg}.

At $g = 0$, the coupling between blocks is completely suppressed; and the operator has pure point spectrum, with eigenvalues given by the union of the spectrum of the individual matrices $V(x)$, $x \in \mathbb{Z}^d$. Let $\lambda_j(x)$, $j = 1, \ldots, N$, denote the eigenvalues associated to the matrix $V(x)$, with corresponding eigenvectors $\mathbf{v}_j(x)$, $j = 1, \ldots, N$, in $\mathbb{C}^N$. Then, for every x, the distribution of the eigenvalues is approximately given by Wigner's semicircle density $(2\pi)^{-1}\sqrt{(4-\lambda^2)_+}$, and the gaps between the eigenvalues (for fixed x) are typically of order N^{-1}.

For positive $g = \frac{a}{\sqrt{N}}$, second-order perturbation theory predicts that the eigenvalues $\lambda_j(x)$ shift by a quantity close to

$$\frac{a^2}{N} \sum_{y \sim x} \sum_{k=1}^{N} \frac{\left|\left\langle \mathbf{v}_k(y), \mathbf{v}_j(x) \right\rangle\right|^2}{\lambda_j(x) - \lambda_k(y)} \;\approx\; \frac{a^2 d}{\pi N}\,\mathrm{P.V.}\int_{-2}^{2} \frac{\sqrt{4-\lambda^2}}{\lambda_j(x) - \lambda}\mathrm{d}\lambda = \frac{a^2 d\;\lambda_j(x)}{N},$$

that is, comparable to the mean gap.

Though the series in a provided by Rayleigh–Schrödinger (infinite order) perturbation theory has zero radius of convergence, the considerations of the previous paragraph provide an indication that the scaling $g = a/\sqrt{N}$ is natural.

Supersymmetric models.

Another perspective on the models (1.2), and random operators in general, is given by dual supersymmetric models, which were introduced by Efetov [21], following earlier work by Wegner and Schäfer [35, 42]; see further the monograph of Wegner [44] and the mathematical review of Spencer [41]. In the supersymmetric approach, $\mathbb{E}|(H-z)^{-1}(x,y)|^2$ is expressed as a two-point correlation in a dual supersymmetric model. Fixing $a > 0$ and setting $g = \frac{a}{\sqrt{N}}$, the supersymmetric models dual to (1.2) should converge, as $N \to \infty$, to a supersymmetric σ-model with $U(1,1 \mid 2)$ symmetry (in the unitary case) and $OSp(2,2 \mid 4)$ symmetry (in the orthogonal case), at temperature determined by a and the value of the density of states. These σ-models are conjectured to exhibit a phase transition in dimension $d \geq 3$. This provides additional support for the scaling (1.3).

As to rigorous results, a mathematical proof of the existence of phase transition for the supersymmetric σ-models remains a major challenge. Progress was made by Disertori, Spencer, and Zirnbauer [19, 20], who rigorously established the existence of phase transition for a supersymmetric σ-model with the simpler $OSp(2 \mid 2)$ symmetry. Presumably, the analysis of the Efetov σ-models presents additional challenges.

The convergence of the dual supersymmetric models to the corresponding σ-models is, to the best of our knowledge, not yet mathematically established. Moreover, a strong form of convergence is required to infer the existence of a phase transition before the limit; convergence of the action does not by itself suffice.

Lower bounds on the density of states.

The upper bound in Theorem 2 is often sharp up to a multiplicative constant. One can obtain complementary bounds to Theorem 2 in terms of the second moment

$$s_2^2 := \frac{1}{\sum_{j=1}^k N_j} \mathbb{E} \operatorname{tr} H^2 = \frac{1}{\sum_{j=1}^k N_j} \mathbb{E} \int \lambda^2 \mathcal{N}(H, \mathrm{d}\lambda) \,.$$

Namely, for any $0 < t < s_2$ there exists an interval $I \subset [-2s_2, 2s_2]$ with $|I| = t$ for which

$$\mathbb{E}\mathcal{N}(H, I) \geq \frac{1}{10 s_2} \sum_{j=1}^k N_j \, |I| \,, \tag{4.1}$$

since the failure of this bound would imply that

$$\mathbb{E}\mathcal{N}(H, [-2s_2, 2s_2]) \leq \left\lceil \frac{4s_2}{t} \right\rceil \frac{t}{10 s_2} \sum_{j=1}^k N_j \leq \frac{1}{2} \sum_{j=1}^k N_j$$

in contradiction to Chebyshev's inequality. In the applications to the orbital models (1.2) and to Gaussian band matrices (e.g., Definition 1.4 with ψ as in (1.15) and ϕ decaying sufficiently fast), the quantity s_2 is itself bounded by a constant, whence Theorem 2 is sharp in these cases.

It is plausible that, for orbital models and band matrices, bounds of the type (4.1) also hold for individual intervals sufficiently close to the origin; see Wegner [43] for the case $N = 1$.

Three open questions.

We use this opportunity to recapitulate a few of the open questions pertaining to the block Anderson and Wegner orbital models (1.2).

(1) Is it true that in dimension $d \geq 3$ one has the following converse to Theorem 1: for $g \geq C(d)N^{-1/2}$, there exist energies λ at which exponential decay of the form (1.6) does not hold, at least, for large N? Absence of exponential decay for an interval of energies could be considered as a signature of delocalisation.

(2) Consider the case of fixed g and $N \to \infty$. Is it true that the density of states, the density of the measure (1.11), converges, as $N \to \infty$, in uniform metric? Convergence in the weak-$*$ metric (to an explicit limiting measure) was proved by Khorunzhiy and Pastur in [27, 31]. To upgrade their result to uniform convergence, it would suffice (by a compactness argument) to show that the density of states is equicontinuous in N as a function of the spectral parameter λ.

(3) Is the density of states a smooth function of the spectral parameter? It is expected to be analytic for all values of $g > 0$ and $N \geq 1$. See further the results of [14, 15] discussed after the proof of Corollary 1.3, and the work [16] and references therein pertaining to the Anderson model ($N = 1$).

Funding

ISF grants [1048/11 and 861/15], IRG grant SPTRF, and the European Research Council start-up grant [678520] (LocalOrder) to R.P.; The Fund For Math and NSF grant [DMS-1500386] to J.S.; ISF grant [147/15 to M.S.]; and European Research Council start-up grant [639305] (SPECTRUM) to S.S.

Acknowledgements

Tom Spencer encouraged us to look for a proof to the estimate (1.17). We had the pleasant opportunity to discuss various aspects of supersymmetric σ-models with him as well as with Yan Fyodorov and Tanya Shcherbina. The current project profited greatly from the helpful advice of Michael Aizenman, and also crucially relies on the joint work [4]. We thank them very much.

References

[1] Abou-Chacra, R., D. J. Thouless, and P. W. Anderson. “A selfconsistent theory of localization.” *J. Phys. C Solid State* 6, no. 10 (1973): 1734.

[2] Aizenman, M. "Localization at weak disorder: some elementary bounds." *Rev. Math. Phys.* 6, no. 5A (1994): 1163–82.

[3] Aizenman, M. and S. Molchanov. "Localization at large disorder and at extreme energies: an elementary derivation." *Comm. Math. Phys.* 157, no. 2 (1993): 245–78.

[4] Aizenman, M., R. Peled, J. Schenker, M. Shamis, and S. Sodin. "Matrix regularizing effects of Gaussian perturbations." *Commun. Contemp. Math.* 19, no. 3 (2017): 1750028, 22 pp.

[5] Aizenman, M., J. Schenker, R. Friedrich, and D. Hundertmark. "Finite-volume fractional-moment criteria for Anderson localization." *Comm. Math. Phys.* 224, no. 1 (2001): 219–53.

[6] Aizenman, M. and S. Warzel. *Random operators. Disorder effects on quantum spectra and dynamics*. Graduate Studies in Mathematics, 168. Providence, RI: American Mathematical Society, 2015. xiv+326 pp.

[7] Akhiezer, N. I. *The Classical Moment Problem and Some Related Questions in Analysis.* New York: Hafner Publishing, 1965. x+253 pp. Translated by N. Kemmer.

[8] Anderson, P. W. "Absence of diffusion in certain random lattices." *Phys. Rev.* 109, no. 5 (1958): 1492.

[9] Bapst, V. "The large connectivity limit of the Anderson model on tree graphs." *J. Math. Phys.* 55, no. 9 (2014): 092101, 20.

[10] Bogachev, L. V., S. A. Molchanov, and L. A. Pastur. "On the density of states of random band matrices." *Mat. Zametki* 50, no. 6 (1991): 31–42, 157 (Russian).

[11] Casati, G., L. Molinari, and F. Izrailev. "Scaling properties of band random matrices." *Phys. Rev. Lett.* 64 (1990): 1851–54.

[12] Casati, G., B. V. Chirikov, I. Guarneri, and F. M. Izrailev. "Band-random-matrix model for quantum localization in conservative systems." *Phys. Rev. E.* 48 (1993): R1613.

[13] Combes, J.-M., F. Germinet, and A. Klein. "Generalized eigenvalue-counting estimates for the Anderson model." *J. Stat. Phys.* 135, no. 2 (2009): 201–16.

[14] Constantinescu, F. "The supersymmetric transfer matrix for linear chains with nondiagonal disorder." *J. Stat. Phys.* 50, no. 5–6 (1988): 1167–77.

[15] Constantinescu, F., G. Felder, K. Gawędzki, and A. Kupiainen. "Analyticity of density of states in a gauge-invariant model for disordered electronic systems." *J. Stat. Phys.* 48, no. 3–4 (1987): 365–91.

[16] Constantinescu, F., J. Fröhlich, and T. Spencer. "Analyticity of the density of states and replica method for random Schrödinger operators on a lattice." *J. Statist. Phys.* 34, no. 3–4 (1984): 571–96.

[17] Disertori, M. and M. Lager. "Density of states for random band matrices in two dimensions." arXiv:1606.09387

[18] Disertori, M., H. Pinson, and T. Spencer. "Density of states for random band matrices." *Comm. Math. Phys.* 232, no.1 (2002): 83–124.

[19] Disertori, M. and T. Spencer. "Anderson localization for a supersymmetric sigma model." *Comm. Math. Phys.* 300, no. 3 (2010): 659–71.

[20] Disertori, M., T. Spencer, and M. R. Zirnbauer. "Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model." *Comm. Math. Phys.* 300, no. 2 (2010): 435–86.

[21] Efetov, K. B. "Supersymmetry and theory of disordered metals." *Adv. Phys.* 32, no.1 (1983): 53–127.

[22] Erdős, L. and A. Knowles. "Quantum diffusion and delocalization for band matrices with general distribution." *Ann. Henri Poincaré* 12, no. 7 (2011): 1227–319.

[23] Erdős, L., H.-T. Yau, and J. Yin. "Bulk universality for generalized Wigner matrices." *Probab. Theory Related Fields* 154, no. 1–2 (2012): 341–407.

[24] Fröhlich, J. and T. Spencer. "Absence of diffusion in the Anderson tight binding model for large disorder or low energy." *Comm. Math. Phys.* 88, no. 2 (1983): 151–84.

[25] Fyodorov, Y. V. and A. D. Mirlin. "Scaling properties of localization in random band matrices: A σ-model approach." *Phys. Rev. Lett.* 67 (1991): 2405–9.

[26] Khorunzhiǐ, A. M. and S. A. Molchanov, L. A. Pastur. "Distribution of the eigenvalues of random band matrices in the limit of their infinite order." *Teoret. Mat. Fiz.* 90, no. 2 (1992): 163–78.

[27] Khorunzhy [Khorunzhiy], A. M., and L. A. Pastur. "Limits of infinite interaction radius, dimensionality and the number of components for random operators with off-diagonal randomness." *Comm. Math. Phys.* 153, no. 3 (1993): 605–46.

[28] Levin, B. Ya. *Lectures on Entire Functions*. In collaboration with and with a preface by Yu. Lyubarskii, M. Sodin and V. Tkachenko. Translated from the Russian manuscript by Tkachenko. Translations of Mathematical Monographs, vol. 150. Providence, RI: American Mathematical Society, 1996. xvi+248 pp.

[29] Minami, N. "Local fluctuation of the spectrum of a multidimensional Anderson tight binding model." *Comm. Math. Phys.* 177, no. 3 (1996): 709–25.

[30] Oppermann, R. and F. Wegner. "Disordered system with n orbitals per site: $1/n$ expansion." *Z. Phys. B Con. Mat.* 34, no. 4 (1979): 327–48.

[31] Pastur, L. "On connections between the theory of random operators and the theory of random matrices." *St. Petersburg Math. J.* 23, no.1 (2012): 117–37.

[32] Pastur, L. and A. Figotin. *Spectra of Random and Almost-Periodic Operators*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 297. Berlin: Springer, 1992. viii+587 pp.

[33] Pastur, L. A. and M. Shcherbina. *Eigenvalue Distribution of Large Random Matrices*, vol. 171. Providence, RI: American Mathematical Society, 2011.

[34] Pchelin, V. "Poisson statistics for random deformed band matrices with power law band width." arXiv:1505.06527.

[35] Schäfer, L. and F. Wegner. "Disordered system with n orbitals per site: Lagrange formulation, hyperbolic symmetry, and Goldstone modes." *Z. Phys. B Con. Mat* 38, no. 2 (1980): 113–26.

[36] Schenker, J. H. "Eigenvector localization for random band matrices with power law bandwidth." *Comm. Math. Phys.* 290 (2009): 1065–97.

[37] Schenker, J. H. "How large is large? Estimating the critical disorder for the Anderson model." *Lett. Math. Phys.* 105, no. 1 (2015): 1–9.

[38] Shcherbina, M. and T. Shcherbina. "Transfer matrix approach to 1d random band matrices: density of states." *J. Stat. Phys.* (2016), arXiv:1603.08476.

[39] Sodin, S. “An estimate for the average spectral measure of random band matrices.” *J. Stat. Phys*. 144, no. 1 (2011): 46–59.

[40] Spencer, T. “Random Banded and Sparse Matrices.” In *Oxford Handbook of Random Matrix Theory*, edited by G. Akemann, J. Baik and P. Di Francesco, Oxford Handbooks in Mathematics. Oxford: Oxford University Press, 2011. xxxii+919 pp.

[41] Spencer, T. “Duality, Statistical Mechanics, and Random Matrices.” In *Current Developments in Mathematics, 2012*, edited by D. Jerison, M. Kisin, T. Mrowka, R. Stanley, H.-T. Yau and S.-T. Yau, Somerville, MA: International Press, 2013. iv+260 pp.

[42] Wegner, F. “Disordered system with n orbitals per site: $n = \infty$ limit.” *Phys. Rev. B* 19, no. 2 (1979): 783.

[43] Wegner, F. “Bounds on the density of states in disordered systems.” *Z. Phys. B* 44, no. 1–2 (1981): 9–15.

[44] Wegner, F. *Supermathematics and its Applications in Statistical Physics. Grassmann Variables and the Method of Supersymmetry.* Lecture Notes in Physics, 920. Heidelberg: Springer, 2016. xvii+374 pp.